REMARKS/ARGUMENTS

These remarks are made in response to the Office Action of October 04, 2004 (Office Action). This response is filed along with a petition for a one month retroactive extension of time and an appropriate fee.

In paragraphs 2-3, of the Office Action, the Examiner has rejected claims 1-4, 6-9, 11-14, 16, 17, 19, 21, 22, 24-28, 34, 36-40, 46-49, 51-54, 56-69, 61 and 62 under 35 U.S.C. § 102(b) as being unpatentable over U.S. Patent No. 5,799,273 to Mitchell, *et al.* (Mitchell). In paragraphs 4-5, the Examiner has rejected claims 5, 10, 15, 18, 20, 23, 29-33, 35, 41-45, 50, 55, 60 and 63 under 35 U.S.C. § 103(a) as being unpatentable over Mitchell in view of U.S. Patent No. 5,680,511 to Baker, *et al.* (Baker).

In response to the Office Action, Applicants have amended claims 1, 19, and 46 to clarify that the model context-enhanced database includes entries of speech segments and associated text segments and to clarify that the input is specifies a context used to anticipate content within a speech signal that is to be converted to text. Claims 2, 7, 8, 47, 52, and 53 have been modified to contain language consistent with the amendments to the independent claims. Claims 3, 4, 22, 34, 48, and 49 have been modified to clarify that the context-enhanced database is searched for matches before another database is searched. Claims 5 and 50 have been modified to clarify that the context-enhanced database is created from the input and from entries within a second database. Claims 9, 24, and 54 have been modified to clarify that input is derived during a pre-processing step based upon active applications. Claims 10 and 55 have been amended to clarify that the extracted content generates a word list. Claims 11 and 56 have been amended to clarify that the word list and a context-independent database are used in creating the context-enhanced database. Claims 16, 17, 61, and 62 have been amended to clarify that the context-enhanced database is updated when the context changes. Claim 20 has been amended to further limit the elements of a Markush grouping. New claims 64-71 have

20

been added to emphasize various inventive arrangements disclosed within the specification, which are largely contained within the aforementioned amendments.

Support for these amendments can be found in FIGS 1-8 (specifically within FIG. 2 and 3), at page 8, lines 3-8 and 17-19, at page 9, lines 10-13, between page 9, line 17 and page 11, line 18 (specifically page 11, lines 1-9), at page 13, lines 10-13 and lines 22-28, and throughout the specification. No new matter has been added as a result of these amendments.

It may be helpful to briefly review the features of Applicants' invention before addressing the rejections on the art. The Applications provide a solution to problems relating to speech-to-text converting utterances when utilizing large speech vocabularies, which can also be called grammars or databases for utterance matching. The solution utilizes a context-enhanced database, which can be formed from a larger context-independent database and contextual input. Whenever a speech-to-text conversion is attempted, an utterance is parsed into a unit that a speech engine is configured to process. The context-enhanced database can be searched for matches for the parsed units. Matches from the context-enhanced database, which is of a smaller size than the larger context-independent database, consume significantly fewer computational resources and generally produce more accurate results than matches performed against the context-independent database. When no matches are found in the context-enhanced database, the context independent database can be utilized to determine a text segment that corresponds to the parsed unit. The present invention further teaches that utterances can be textually edited by a user.

In one embodiment, the Applicants' solution can include a pre-processing step that defines content for an anticipated voice-generated output that is to be generated by a user of a computer system upon which the method of speech recognition executes. During the pre-processing step, input can be automatically deriving based upon active applications

21

currently executing upon the computer system. For example, content from electronic documents enabled within the active applications can be extracted to generate a word list. The derived input can include the word list. Moreover, the context dependent database can be formed from those entries of the context-independent database having words included within the word list. Additionally, changes in the active applications can be automatically detected, resulting in new input being derived, and the context-dependant database being updated.

Turning the rejections on the art, claims 1-4, 6-9, 11-14, 16, 17, 19, 21, 22, 24-28, 34, 36-40, 46-49, 51-54, 56-69, 61 and 62 have been rejected under 35 U.S.C. § 102(b) as being anticipated by Mitchell. Mitchell teaches a word processor application that incorporates speech-to-text generated output into an editable textual document. Linkages between speech utterances and text-to-speech generated output are maintained. When textual edits to the document are made, corresponding and appropriate changes to the links are made. Hence, the editable text document permits a speech-to-text generated document to be proofread and for a user to make suitable changes and/or corrections. When a document is saved, the links can be used to determine whether speech misrecognitions occurred based upon the edits. Determined misrecognitions and their corrections can be used to train a speech recognition system to improve the accuracy of the system.

Referring to claims 1 and 46, the Applicants claim the steps of:

(a)   receiving an input that specifies a context in which the speech recognition system processes speech such that the speech recognition system is able to anticipate content within a speech signal to be received based upon the context;

(b)   automatically creating a context-enhanced database using information derived from said input;

22

{FT255007;1}

(c)   preparing a first textual output from a speech signal by performing a speech recognition task to convert said speech signal into said first textual output comprising computer-processable text segments, wherein said context-enhanced database is accessed to improve the speech recognition rate, and wherein selective ones of the computer processable text segments are generated by matching a portion of said speech signal against an entry within the context-enhanced database, said context-enhanced database including a plurality of entries, each entry comprising a speech utterance and a corresponding textual segment for the speech utterance;

(d)   enabling editing of said first textual output to generate a final textual output; and

(e)   making said final textual output available.

It is clear from the claims that the context-enhanced database is an acoustic speech-processing grammar that generates text output by matching received speech utterances against entries in a database table, as specifically noted at page 11, lines 1-8. The speech recognition rate is improved since the context-enhanced database include fewer entries than a conventionally utilized context-independent database, and since the number of entries in an acoustic speech processing grammar dramatically affects the speech recognition rate, as noted at page 3, lines 3-16. Mitchell fails to provide any corresponding teaching, and instead focuses upon modifying a contextual model and not an acoustic model, as noted at column 8, lines 9-22, in order to update a user's model based upon edits. Consequently, Mitchell teaches a speech recognition system that is tailored for a particular user, which is a type of conventional system described at page 3, lines 4-11 of the Applicants' background.

Mitchell fails to construct a context enhanced database based upon an input used by the speech recognition system to anticipate content within a speech signal to be

received based upon a determined context. Mitchell does not teach that content of speech to be received is to be anticipated in any fashion.

Further, Mitchell improves a contextual model based upon text edit to improve speech recognition accuracy. Mitchell does not teach a technique that improves an acoustic model, as claimed by the Applicants. Mitchell references an acoustic model and operations that are to be performed against it at FIG. 5, steps S10 and S11, as noted by column 8, lines 14-18, and at FIG 8B step S83.

The majority of the Examiner provided citations from Mitchell are specifically directed to teachings for updating a contextual model, which are not equivalent teachings. In the field of speech processing, different techniques, algorithms, data structures, and the like are used for performing acoustic matches and for improving accuracy of speech-converted-text using grammatical rules. The difference is that the acoustic model translates digitally encoded sounds into textual equivalents and the contextual model of Mitchell modifies speech produced textual output in accordance with structural linguistic rules on how words are to be put together to form meaningful combinations defined by a particular language. Applicants teach the pre-processing (before the acoustic model is used) of available input to generate a context-dependent acoustic model, referred to by the applicants as a context-enhanced database.

Accordingly, the Examiner cited references from Mitchell providing teachings to update a context model fail to teach or suggest the Applicants' claimed limitations pertaining to an acoustic model. These references include: column 8, lines 9-12, FIG. 5, Step 12, and FIG. 15, Step 163.

In order for a 35 U.S.C. § 102 to be validly asserted against a claim, the cited reference must expressly or implicitly include each and every claimed limitation. Mitchell fails in this regard. Specifically, Mitchell fails to expressly or implicitly teach the receiving of input that specifies a context used to anticipate yet to be received input,

24

the creating of a context-enhanced database from the received input, and fails to teach utilizing the context-enhanced database as an acoustic model to generate text from speech utterances.   Accordingly, since Mitchell fails to explicitly or implicitly teach each claimed limitation, the 35 U.S.C. § 102(e) rejections to claims 1 and 46, as well as dependent claims 2-4, 6-9, 11-14, 16-17, 47-49, 51-54, 56-59, and 61-62 should be withdrawn, which action is respectfully requested.

Additionally, independent claim 19 contains these same limitations not expressly or implicitly taught by Mitchell.   Applications, therefore, respectfully request the 35 U.S.C. § 102 rejections to claim 19 and dependent claims 21-22, 24-28, 34, 36-40 also be withdrawn.

In paragraphs 4-5, the Examiner has rejected claims 5, 10, 15, 18, 20, 23, 29-33, 35, 41-45, 50, 55, 60 and 63 under 35 U.S.C. § 103(a) as being unpatentable over Mitchell in view of Baker.   Baker teaches a speech recognition method for training a speech recognition system to recognize new words without expressly training a model by prompting a user for each word in a user-specific language model.   Baker is to be used in the context of unrecognized words or activated when a word is not found in a language model, as noted by column 5, lines 8-24.   Baker teaches a methodology that provides user-selectable options for the unknown word to make the training process less cumbersome.

The Examiner cites Baker to cure perceived deficiencies with Mitchell's contextual model.   As previously noted, the techniques taught for the contextual model of Mitchell does not translate into teachings applicable for an acoustic model, the subject of the Applicants claims.

Further, there is no motivation present within Mitchell and Baker to combine the references for 35 U.S.C. § 103 purposes based upon the Applicants' existing claims.   That is, Mitchell teaches a system and method for training speech processing vocabularies.

Baker teaches a solution for training speech vocabularies for new words. Hence, Baker may provide teachings applicable to Mitchell to enhance the training of a user-specific speech vocabulary, which is not pertinent to the Applicants claims.

That is, the teachings of Mitchell and Baker should not be combined for 35 U.S.C. § 103 purposes as the only motivation to attempt the suggested combinations is provided by the Applicants own disclosure. Attempts to combine Mitchell and Baker would thus be motivated only by hindsight logic of the Applicants' invention and not based upon the references themselves or teachings included in these references. Moreover, neither Mitchell nor Baker identity the problems with speech processing utterances using large speech vocabularies, as noted at page 3, lines 4-16 of the Applicant's background, and instead teach methodologies that would further increase the size of a vocabulary without alleviating misrecognition and recognition rate problems that accompany the growth of a speech recognition vocabulary. Consequently, neither Mitchell nor Baker even recognize the problem solved by the Applicants' claimed invention, nor do Mitchell or Baker teach or suggest the Applicants claimed solution to this identified problem.

Specifically, Baker fails to teach or suggest the receiving of input that specifies a context used to anticipate yet to be received input that is to be speech-converted into text, the creating of a context-enhanced database from the received input, and fails to teach utilizing the context-enhanced database as an acoustic model to generate text from speech utterances. In this regard, Baker fails to cure the deficiencies of Mitchell, which also fails to teach or suggest the above claimed limitations.

Since under 35 U.S.C. § 103 each claimed limitation must be taught by or be legally obvious in light of a cited reference or combination of references, and since each claimed limitation of independent claims 1, 19, and 46 are not taught or suggested by Mitchell, Baker, or combinations thereof, the 35 U.S.C. § 103 rejections to claims

Appln. No. 09/910,657
Amdt. dated Jan. 11, 2005
Reply to Office Action of Oct. 04, 2004
Docket No. CH9-2000-0004 (246)

dependent upon claims 1, 19, and 46, which include claims 5, 10, 15, 18, 20, 23, 29-33, 35, 41-45, 50, 55, 60 and 63 should be removed, which action is respectfully requested.

The Applicants believe that this application is now in full condition for allowance, which action is respectfully requested. The Applicants request that the Examiner call the undersigned if clarification is needed on any matter within this Amendment, or if the Examiner believes a telephone interview would expedite the prosecution of the subject application to completion.

Respectfully submitted,

Date: _11 Jan 2005_　　　　　　　　　　_____

Gregory A. Nelson, Registration No. 30,577
Richard A. Hinson, Registration No. 47,652
Brian K. Buchheit, Registration No. 52,667
AKERMAN SENTERFITT
Customer No. 40987
Post Office Box 3188
West Palm Beach, FL 33402-3188
Telephone: (561) 653-5000

27

(FT255007;1)